Public-Key-Kryptosysteme und Arithmetische Geometrie

Frey, Gerhard

Veröffentlicht in:
Jahrbuch 1996 der Braunschweigischen Wissenschaftlichen Gesellschaft, S.207-221

Verlag Erich Goltze KG, Göttingen

GERHARD FREY

Public-Key-Kryptosysteme und Arithmetische Geometrie

Institut für Experimentelle Mathematik
Universität Essen

Herr Präsident, sehr geehrte Damen und Herren!

Ich möchte mich zunächst ganz herzlich für die große Ehre bedanken, die die Verleihung der Gauß-Medaille für mich darstellt.

In der sehr freundlichen Begründung beziehen Sie sich auf die Lösung des Fermatschen Problems, das zweifellos zu den berühmtesten Fragen der Mathematik gehört, und tatsächlich haben wir Mathematiker allen Grund zu feiern: Der (weitgehende) Beweis der Taniyama-Vermutung für elliptische Kurven durch Andrew Wiles ist ein Jahrhundertereignis in unserer Wissenschaft; daß für ihn dazu die Fermatsche Vermutung ein beflügelnder Antrieb war und daß ich ein wenig zu dieser Motivierung beitragen konnte, erfüllt mich mit Freude.

Ich bin sicher, daß dieser Beweis der Fermatschen Vermutung auch Gauß gefallen hätte: Nicht Zufälligkeiten oder unzählige Fallunterscheidungen, sondern tiefe, strukturelle Aussagen über elliptische Kurven führen zum Ziel, und diese Aussagen wiederum beruhen auf arithmetischen Eigenschaften der Galoisgruppe der rationalen Zahlen, die regiert werden von der Theorie der Modulformen, wie es die viel kühnere und allgemeine "Langlands-Philosophie" voraussagt.

Der Wiles'sche Beweis der Taniyama-Vermutung und damit auch der Fermat-Vermutung bestätigt eindrucksvoll die Kraft der uns zur Verfügung stehenden Methoden der Arithmetischen Geometrie, die Zahlentheorie, Geometrie und Funktionentheorie zu einem fruchtbaren Zusammenspiel bringt; daß diese Theorie sich immer weitere Anwendungsfelder auch im "praktischen Leben" erschließt, hätte Gauß vielleicht überrascht, aber sicher nicht befremdet, ist er doch immer der Verbindung von Mathematik und ihren Anwendungsmöglichkeiten aufgeschlossen gegenübergestanden.

Über eine dieser Anwendungsmöglichkeiten *möchte* ich jetzt berichten. Sie beruht auf der *Digitalisierung* der Nachrichtenübertragung, und diese, von

[1]Ausarbeitung eines Vortrags anläßlich der Verleihung der Gauß-Medaille 1996

Ingenieuren getroffene Entscheidung, öffnet die Theorie und Praxis der Datenübertragung für mathematische Methoden aus der *diskreten* Mathematik, also auch für Methoden aus der Arithmetischen Geometrie. Ein besonders fruchtbares Feld ist dabei die *Datensicherheit*, in das wir unsere Kenntnisse aus der arithmetischen Geometrie, z.B. über elliptische Kurven und Modulkurven, einbringen wollen.

1 Public-key-Kryptosysteme und Diskrete Logarithmen

Die Sicherheit bei der Übertragung und Speicherung von Daten setzt zunächst Zuverlässigkeit der verwendeten Geräte und Netze voraus, sie verlangt das Vermeiden von Übertragungsfehlern, und sie fordert schließlich die Nichtmanipulierbarkeit der Daten, z.B. sollen Verfälschungen von Inhalt, Absender, Empfänger und nichtautorisierte Weitergabe von Daten ausgeschlossen werden. Darüber hinaus soll die Kommunikation aber offen, unkompliziert und vor allem schnell und billig sein.

Die Verwirklichung dieser Ziele ist nur erreichbar, wenn schon bei dem Entwurf der Kommunikationssysteme im Hard- und Softwarebereich, ausgehend von den Anforderungen und spezifischen Eigenarten der Anwender, eng zwischen Ingenieuren (die z.B. für die Zuverlässigkeit kompetent sind), Informatikern und Mathematikern zusammengearbeitet wird. Dabei können die Mathematiker bei der Entwicklung schneller Algorithmen (z.B. für die Arithmetik sehr großer Zahlen), bei der Fehlervermeidung durch Codierungstheorie und eben auch bei der Fälschungs- und Abhörsicherheit durch den Einsatz von kryptographischen Methoden wichtige Beiträge liefern.

Ich möchte jetzt auf den zuletztgenannten Aspekt näher eingehen. Grundsätzlich stellt die Kryptographie Algorithmen bereit, die allgemein lesbare Daten zu "verschlüsselten" Daten transformieren, die dann nur noch von autorisierten Personen "entschlüsselt" und damit wieder lesbar gemacht werden können. Seit langem sind dazu sogenannte symmetrische Verfahren entwickelt worden, die darauf beruhen, daß die Kommunikationspartner *gemeinsam* ein Verschlüsselungs- und Entschlüsselungsverfahren (das z.B. Methoden aus der Kombinatorik oder Gruppentheorie verwendet) besitzen und benützen, das sonst niemandem bekannt ist.

Diese Verfahren arbeiten schnell, einfach und sind sehr sicher, falls - und das ist das große Problem - die Schlüssel wirklich geheim bleiben. Dies verursacht vor allem bei der Schlüsselübergabe Probleme, insbesonders, wenn viele Kommunikationspartner betroffen sind.

Sehr elegant wird diese Aufgabe von Public-key-Kryptoverfahren umgangen: Es gibt pro Partner zwei Schlüssel: Einer ist geheim und nur ihm bekannt, ein

zweiter ist öffentlich; die Kommunikation geschieht mit Hilfe der öffentlichen Schlüssel, die Auswertung benutzt den geheimen Schlüssel.

Die bekannten Public-key-Verfahren haben gegenüber symmetrischen Verfahren den Nachteil, langsamer zu sein. Ihr Anwendungsbereich ist dementsprechend zu begrenzen: Sie können zum *Schlüsselaustausch* für symmetrische Verfahren verwendet werden, vor allem aber können sie zu *Authentifikationen* (von Absender, Inhalt, Adressat) von Botschaften verwendet werden. Dieses Authentifizieren wird zunehmend an Bedeutung gewinnen, z.B. bei der Erteilung von Zugangsberechtigungen zu Netzen und Datenbanken, bei "elektronischen Unterschriften", beim Abschluß von Versicherungen, bei der Schadensregulierung $\cdots$.

Grundlegend für alle Public-key-Verfahren ist das Konzept von "Falltürfunktionen": Dies sind Abbildungen, die sehr schnell ausgewertet werden können, bei denen aber die Umkehrabbildung zwar existiert, jedoch praktisch nicht berechnet werden kann.

Es werden also Funktionen

$$f : A \longrightarrow B$$

gesucht, für die für alle $a \in A$ der Wert $f(a) = b$ schnell berechnet werden kann, für die aber für gegebenes b das Argument a, für das $f(a) = b$ gilt, nur sehr schwer bestimmt werden kann.

Realistische Forderungen sind: Die Berechnung von $f(a)$ benötigt auf einem gängigen PC etwa 100 ms, die Umkehrung dauert, unter Verwendung aller verfügbaren Rechenressourcen der Welt, mindestens 1000 Jahre!

Es ist klar, daß die erste Bedingung aus der Konstruktion zwingend folgen muß, sehr schwierig ist es, die zweite Bedingung zu gewährleisten. Man kann immer nur Aussagen über die Sicherheit des verwendeten Systems "nach bestem Wissen und Gewissen", basierend auf dem heutigen Kenntnisstand und der jetzigen Rechnertechnologie machen, und deshalb sollte das gewählte Kryptosystem so flexibel sein, daß es ohne großen Aufwand verändert werden kann.

Ein Beispiel dazu: Ein weit verbreitetes Verfahren, das sogenannte RSA-Verfahren, benutzt für f eine Funktion, deren Umkehrung die Faktorisierung von Zahlen der Form $n = p_1 \cdot p_2$ erfordert, wobei n bekannt und p_1 und p_2 unbekannt sind. Dabei sind p_1 und p_2 etwa gleich große Primzahlen, also ungefähr gleich $\sqrt{n}$. Die naive Methode würde $\sqrt{n}$ Versuche zur Faktorisierung benötigen, und für $n \approx 10^{40}$ wäre damit unsere Sicherheitsbedingung erfüllt: Es sind probabilistisch 10^{20} Versuche zur Berechnung von Urbildern von f notwendig.

In den letzten Jahrzehnten wurden, unter Benutzung von Ergebnissen der algebraischen Zahlentheorie und auch der arithmetischen Geometrie, aber immer effektivere Methoden für das Faktorisieren von Zahlen gefunden. Für Kenner seien die Stichworte "Algebraic number sieve" und "Index-Calculus-Verfahren" genannt. Deshalb werden heute beim RSA-Verfahren Zahlen n mit über 300 Dezimalstellen (1024 bits) verwendet, und es ist abzusehen, daß bald auf 2048 bits gegangen werden muß. Dies macht natürlich sowohl die Erzeugung als auch die Verwaltung der Schlüssel (das sind die geheimen Primzahlen p_1 und p_2) immer aufwendiger, außerdem wird durch die Größe der Zahlen die Schnelligkeit des Verfahrens beeinträchtigt.

Diese Probleme treten gegenwärtig bei dem Verfahren, das ich jetzt vorstellen will, noch nicht auf.

Die erste Beobachtung, die wir benötigen, ist: In jeder Gruppe A, in der man die Gruppenaddition $+$ überhaupt berechnen kann, ist die Vielfachenbildung schnell: Wählen wir $P_0 \in A, n \in \mathbb{N}$ und definieren wir

$$f_{P_0}(n) := n \cdot P_0,$$

so ist f_{P_0} durch höchstens $2\log_2(n)$ Additionen in A berechenbar. Dies sieht man ein; indem man z.B. n in 2-adischer Ziffernentwicklung in der Form

$$n = \sum_{i=0}^{\log_2(n)} \varepsilon_i 2^i \text{ mit } \varepsilon_i \in \{0,1\}$$

darstellt und beachtet, daß

$$n \cdot P_0 = \sum_{\substack{i=0 \\ \varepsilon_i \neq 0}} P_i$$

mit $P_{i+1} = 2P_i$ ist.

Damit können wir auch leicht einen *Public-key-Schlüsselaustausch* zwischen zwei Partnern C_1 und C_2 organisieren: Öffentlich werden die Gruppe A und das Element P_0 bekannt gemacht. Geheim halten C_1 die Zahl k_1 und C_2 die Zahl k_2, wiederum öffentlich sendet C_1 das Gruppenelement $P_1 = k_1 \cdot P_0$ an C_2, C_2 das Gruppenelement $P_2 = k_2 \cdot P_0$ an C_1. Wenn nun C_1 den Punkt P_2 mit seiner Schlüsselzahl k_1 multipliziert und C_2 das Entsprechende mit P_1 macht, haben beide dasselbe Element $P = k_1 \cdot k_2 \cdot P_0$ erhalten, sie haben also einen gemeinsamen Schlüssel.

Da Außenstehende P_0, P_1 und P_2 kennen, hängt die Sicherheit dieses Schlüsselaustausches in offensichtlicher Weise von der Güte der Funktion f_{P_0} als

Falltür-Funktion ab: Es sollte unmöglich sein, z.B. aus P_1 die Zahl k_1 zu bestimmen.

Aus theoretischen und rechentechnischen Gründen ist es sinnvoll, A als zyklische Gruppe von Primzahlordnung p zu wählen, also

$$A = \{k \cdot P_0; 0 \leq k \leq p-1\};$$

die Zahlen k_1 und k_2 können dann in dem Bereich zwischen 1 und $p-1$ gewählt werden (und sind auch nur modulo p bestimmt). Eine (naive) Methode zur Bestimmung von k_i ist natürlich das Durchprobieren, man muß damit rechnen, ungefähr p Versuche durchschnittlich machen zu müssen. Dies scheint darauf hinzudeuten, daß (da wir 10^{20} Versuche gegenwärtig für undurchführbar halten) p in der Größenordnung 10^{20} gewählt werden kann. Leider gibt es aber ein potentiell in allen Gruppen anwendbares Verfahren von Pollard, das in jeder abelschen Gruppe A der Ordnung p die Berechnung der Zahl k_i probabilistisch auf $\sqrt{p}$ Schritte reduziert, falls es gelingt, A in drei Gebiete A_1, A_2, A_3 so aufzuteilen, daß die Folge

$$P_1 = P; \text{und für } i \geq 2 \quad P_i := \begin{cases} 2P_{i-1} & ; \quad P_{i-1} \in A_1 \\ P + P_{i-1} & ; \quad P_{i-1} \in A_2 \\ P_0 + P_{i-1} & ; \quad P_{i-1} \in A_3 \end{cases}$$

sich auf die Gebiete A_1, A_2, A_3 gleichmäßig verteilt.

Schon die Möglichkeit, daß das Pollardsche Verfahren implementiert werden könnte, zwingt nun dazu, $p \approx 10^{40}$ zu wählen, wenn wir 10^{20} Versuche zur Berechnung von k_i probalistisch erzwingen wollen.

Kommen wir jetzt zu konkreten Gruppen: Die einfachste Art, eine zyklische Gruppe der Ordnung p zu realisieren, ist, in den ganzen Zahlen $\mathbb{Z}$ "modulo p" zu rechnen:

Der ganzen Zahl z wird ihr (kleinster positiver) Rest $r_z \in \{0, \ldots, p-1\}$, der beim Teilen durch p entsteht, zugeordnet, zwei Zahlen z_1 und z_2 heißen äquivalent, wenn $r_{z_1} = r_{z_2}$ (d.h.: $z_1 - z_2$ wird von p), mit $\mathbb{Z}/p\mathbb{Z}$ wird die Menge dieser Äquivalenzklassen bezeichnet, die in eineindeutiger Weise den Resten entspricht, also aus p Elementen besteht. Zwei Elemente $\bar{z_1}, \bar{z_2}$ aus $\mathbb{Z}/p\mathbb{Z}$ werden addiert, indem man die entsprechenden Reste in $\mathbb{Z}$ addiert und vom Ergebnis wieder den Rest nimmt. Es ist leicht zu sehen, daß mit dieser Verknüpfung $\mathbb{Z}/p\mathbb{Z}$ zu einer zyklischen Gruppe der Ordnung p wird, als Erzeuger kann die Klasse jeden Restes ungleich 0, etwa r_0, genommen werden. Unsere Funktion f_{r_0} ist dann gegeben durch $f_{r_0}(k) = r'_k$, mit einer Zahl r'_k zwischen 1 und $p-1$, so daß $kr_0 - r'_k$ durch p teilbar ist. Dieses

r_k' ist sehr schnell zu berechnen. Wie steht es aber mit der oben diskutierten Sicherheitsfrage?

Für einen "Lauscher" ist folgende Aufgabe zu lösen: Zu bekannten Zahlen r_0 und r_k' berechne k, so daß $k \cdot r_0 - r_k'$ durch p teilbar ist. Dies ist sicher leicht machbar, wenn wir ein $\lambda \in \mathbb{Z}$ finden, so daß $\lambda \cdot r_0 - 1$ durch p teilbar ist: k ist der Rest der Zahl $\lambda \cdot r_k'$. Da nach Voraussetzung r_0 und p teilerfremd sind, besagt ein grundlegender Satz der elementaren Zahlentheorie, daß man λ und μ in $\mathbb{Z}$ so findet, daß $\lambda r_0 + \mu p = 1$ ist, und daß zur Bestimmung der Zahlen λ und μ der *Euklidische Algorithmus* verwendet werden kann, der das Ergebnis ungefähr in $\log p$ Schritten liefert. Damit ist unsere Funktion f_{r_0} für kryptographische Zwecke ungeeignet. Immerhin gewinnen wir aber die Erkenntnis, daß $\mathbb{Z}/p\mathbb{Z}$ nicht nur eine additive Gruppe, sondern ein Körper ist (d.h. man kann multiplizieren und zu Elementen ungleich 0 Inverse finden), dessen Elemente leicht in Computern darstellbar sind und in dem man schnell rechnen kann. Es ist nicht schwer, diese effiziente Arithmetik auf beliebige endliche Körper K auszudehnen.

Dies nützen wir aus, um die Präsentation von zyklischen Gruppen von Primzahlordnung für unsere Zwecke geeigneter zu machen: Man weiß, daß die Anzahl der Elemente in endlichen Körpern eine Primzahlpotenz, etwa l^f, ist, und daß die Elemente ungleich 0 eine zyklische Gruppe der Ordnung $l^f - 1$ bilden, wobei jetzt die Verknüpfung die Körpermultiplikation ist. Ist also p ein Teiler von $l^f - 1$, so gibt es ein Element ζ_p aus K mit $\zeta_p{}^p = 1$ und $\zeta_p{}^k \neq 1$ für $1 \leq k \leq p-1$.

Nehmen wir $P_0 = \zeta_p$ und $f_{\zeta_p}(k) = \zeta_p{}^k$, so erhalten wir tatsächlich eine Funktion mit Werten von K^*, die es verdient, "kryptographisch" untersucht zu werden. Die Aufgabe, die der Angreifer nun hat, ist, bei bekanntem Körper K, bekanntem ζ_p und $\zeta_p{}^k$ die Hochzahl k zu bestimmen, also den "Logarithmus zur Basis ζ_p" zu berechnen.

Daher erklärt sich der Name "Diskreter Logarithmus (DL)", der für die ganze Klasse von Falltürfunktionen, über die ich spreche, verwendet wird.

Der wesentliche Schutz vor der Berechnung dieses diskreten Logarithmus besteht darin, daß keine nichttriviale topologische Struktur und damit keine "Analysis" zur Verfügung steht. Angriffe verfolgen deshalb oft die Strategie der *Liftung*, die die Aufgabe der Berechnung des DL in Rechnungen über Körper mit mehr Struktur wie l-adische Körper oder algebraische Erweiterungskörper von $\mathbb{Q}$ von kleinem Grad, zurückführt, und tatsächlich kann man mit Methoden, die denen bei der Faktorisierung großer Zahlen ähneln, die oben beschriebene Funktion f_{ζ_p} relativ schnell ("subexponentiell") umkehren. Dies zwingt uns wie bei den RSA-Verfahren, bei Verwendung von

f_{ζ_p} die Primzahl p sehr groß (z.B. 500-stellig) zu wählen, was natürlich den Rechenaufwand in dem zugehörigen Körper unangemessen in die Höhe treibt.

Die Idee des "Diskreten Logarithmus" ist aber so elegant, daß man sie nicht ohne weiteres aufgeben möchte. Zunächst beobachtet man, daß man die von ζ_p erzeugte Gruppe auf natürliche Weise geometrisch interpretieren kann: Elemente in einem Körper K, die ungleich 0 sind, entsprechen eineindeutig den Punkten auf der Hyperbel G_m

$$XY = 1,$$

deren Koordinaten in K liegen.
Jedem $x \in K\backslash\{0\}$ wird der Punkt $(x, x^{-1}) \in G_m(K)$ zugeordnet. Die Multiplikation in K kann durch Polynome in den Koordinaten (X, Y) beschrieben werden: Für $(x_1, y_1) \in G_m(K), (x_2, y_2) \in G_m(K)$ ist $(x_1, y_1) \otimes (x_2, y_2) = (x_1x_2, y_1y_2)$, und falls $\zeta_p \in K$ ist, so ist die von den Potenzen von ζ_p erzeugte Gruppe gerade gleich $G_m(K)_p$, die Gruppe der "Punkte der Ordnung p" von G_m.

Diese Übersetzung mag künstlich erscheinen. Sie liefert aber das einfachste Beispiel für *Kommutative Algebraische Gruppen* $\mathcal{A}$, die über K definiert sind: $\mathcal{A}$ ist ein geometrisches Objekt (eine *Varietät,* das durch Polynomgleichungen mit Koeffizienten aus K definiert ist). Für Erweiterungskörper L von K sollte man sich die L-rationalen Punkte $\mathcal{A}(L)$ als Elemente $(a_1, \ldots, a_n) \in L^n$ mit geeignetem n, deren Koordinaten die $\mathcal{A}$ beschreibenden Gleichungen erfüllen, vorstellen. Auf dieser Menge $\mathcal{A}(L)$ ist eine Verknüpfung $\oplus$ definiert, die wiederum durch Polynomgleichungen mit Koeffizienten aus K gegeben wird, und die $\mathcal{A}(\mathcal{L})$ zu einer kommutativen Gruppe macht. Für natürliche Zahlen n wird mit $\mathcal{A}(L)_n$ die Untergruppe, bestehend aus den Punkten P aus $\mathcal{A}(L)$, für die $n \cdot P = P \oplus \ldots \oplus P$ (n-fache Addition) das neutrale Element 0 ist, bezeichnet. Die Verallgemeinerung der Funktion f_{ζ_p} ist offensichtlich: Wir nehmen an, daß es in $\mathcal{A}(L)$ einen Punkt $P_0 \neq 0$ gibt, für den für die Primzahl p gilt: $p \cdot P_0 = 0$, und setzen für $1 \leq k \leq p-1$:

$$f_{P_0}(k) = k \cdot P_0$$

Man beachte, daß $k \cdot P_0$ durch ein n-tupel von Elementen aus L gegeben wird, also mit dem n-fachen des Aufwands, mit der Elemente aus L beschreibbar sind, abgespeichert werden kann. Daraus folgt schon, daß man versuchen muß, sowohl n klein als auch den Körper L für Computer gut beschreibbar zu wählen. Zusätzlich ist für Schnelligkeit der Berechnung von f_{P_0} der Grad der Polynome, die $\mathcal{A}$ und $\oplus$ beschreiben, von großer Wichtigkeit, er sollte ebenfalls möglichst klein sein. Schließlich sollten aus Sicherheitsgründen "Liftungen" z.B. im Erweiterungskörper von $\mathbb{Q}$ von kleinem Grad oder in

l-adische Körper, die zur Berechnung der Umkehrfunktion von f_{P_0} verwendet werden können, erschwert werden. Offensichtlich haben wir mit diesem Ansatz zur Konstruktion von Falltürfunktionen den Bereich der Elementaren Zahlentheorie verlassen und kommen nicht umhin, schon zu seiner Beschreibung mehr Mathematik zu verwenden. Deshalb wird der nächste Abschnitt einige Vertrautheit mit Methoden der Algebra und Geometrie voraussetzen müssen; wir ziehen schon hier das

Resümee: Es zeigt sich, daß diese Forderungen, nach dem gegenwärtigen Stand unseres Wissens, erfüllbar sind, wenn wir geeignete *Abelsche Varietäten* über endlichen Körpern K (z.B. über $\mathbb{Z}/p\mathbb{Z}$) von nicht zu großer Dimension wählen und bei dieser Wahl einige leicht einzuhaltende Vorsichtsmaßnahmen beachten, und es gelingt, solche geeignete Varietäten explizit und mit vertretbarem Rechenaufwand zu konstruieren.

2 Kryptographisch geeignete Jacobische Varietäten

Den Hintergrund der jetzt zu beschreibenden Konstruktion von Falltürfunktionen bildet die Theorie der *Abelschen Varietäten* über endlichen Körpern. Folgende Probleme sind zu lösen:

1.) Konstruiere Abelsche Varietäten $\mathcal{A}$ über endlichen Körpern K, die einen K-rationalen Punkt P_0 der Ordnung p besitzen, und

2.) beschreibe diese Punkte und die Vielfachenbildung so, daß möglichst wenig Speicherplatz und Rechenzeit bei der Berechnung von f_{P_0} benötigt wird.

Die Anzahl der Punkte in $\mathcal{A}(K)$ läßt sich dank des Satzes von A. Weil, der dem geometrischen Analogon der Riemann'schen Vermutung entspricht, abschätzen: Falls K q Elemente besitzt, so ist

$$\#\mathcal{A}(K) \approx q^{Dim\,\mathcal{A}},$$

wobei $Dim(\mathcal{A})$ die Dimension von $\mathcal{A}$ als algebraische Varietät (d.h. die Anzahl der algebraisch-unabhängigen Variablen in den Gleichungen, die $\mathcal{A}$ beschreiben) ist. Wir können damit die Forderung 1.) konkretisieren und verlangen, daß p ein Teiler von $\#\mathcal{A}(K)$ ist, und daß $q^{Dim\,\mathcal{A}}$ sehr nahe bei p liegt. (In der Praxis soll $\frac{q^{Dim\,\mathcal{A}}}{p} \leq 10^5$ sein.)
Es empfiehlt sich, die betrachtete Klasse von Abelschen Varietäten auf ganz bestimmte und leichter zu beschreibende Varietäten zu beschränken:
$\mathcal{A}$ wird die *Jacobische Varietät* einer *projektiven Kurve* C sein.
Für uns genügt es, *ebene* Kurven zu betrachten: Es gibt ein homogenes Polynom $\tilde{f}(X, Y, Z)$ mit Koeffizienten aus K, das C beschreibt: Ein Punkt P im 2-dimensionalen projektiven Raum mit den homogenen Koordinaten (x, y, z) ist ein Punkt von C, falls $\tilde{f}(x, y, z)$ gleich 0 ist.

Sehr oft geht man zu *affinen* Teilen von C über. Man betrachtet zunächst alle Punkte auf C, für die $z = 0$ ist. (Diese bilden, so kann man annehmen, eine endliche Menge $\{P_\infty, \ldots, P_{\infty,d}\}$.) Alle anderen Punkte von C können mit *affinen* Koordinaten (x, y), die Nullstellen des Polynoms $\tilde{f}(X, Y) = f(X, Y, 1)$ sind, beschrieben werden. Man nennt $f(X, Y)$ eine affine Gleichung für C.

Sei $\bar{K}$ der algebraische Abschluß von K. Dies ist ein Erweiterungskörper von K, in dem jedes Polynom eine Nullstelle besitzt und der aus Elementen besteht, die Nullstellen von geeigneten Polynomen mit Koeffizienten aus K sind. Die (absolute) *Galoisgruppe* G_K von K besteht aus den Körperautomorphismen von $\bar{K}$, die Elemente aus K festlassen. Die algebraischen Punkte von C sind

$$C(\bar{K}) = \{P_{\infty,1}, \cdots, P_{\infty,d}\} \cup \{(x, y) \in \bar{K} \times \bar{K};\ f(x, y) = 0\}.$$

Die Gruppe G_K operiert auf $C(\bar{K})$ koordinatenweise.

Der Kurve C wird eine wichtige Invariante, ihr *Geschlecht* g, zugeordnet. Falls $g = 0$, ist C ein Kegelschnitt, also eng verwandt mit G_m und damit für uns uninteressant. Wir setzen deshalb ab jetzt voraus, daß $g \geq 1$ ist.

Der Schlüssel zur Beschreibung der Punkte auf der C zugeordneten Jacobischen Varietät J_C und der Addition dieser Punkte wird durch den Satz von *Riemann-Roch* geliefert, der die Theorie der algebraischen Kurven regiert:

Die Menge $J_C(K)$ der K-rationalen Punkte von J_C entspricht den ungeordneten g-tupeln $(P_1, \cdots, P_g)$ von Punkten aus $C(\bar{K})$, die K-rational sind, die also von Elementen aus G_K in sich übergeführt werden.

Zur Beschreibung der Addition setzen wir voraus, daß C einen Punkt mit Koordinaten in K besitzt, von dem wir annehmen dürfen, daß er gleich $P_{\infty,1} =: P_\infty$ ist.

Für Punkte $\mathcal{P}_1 = (P_1^1, \cdots, P_g^1), \mathcal{P}_2 = (P_1^2, \cdots, P_g^2)$ aus $J_C(K)$ ist

$$\mathcal{P}_3 = \mathcal{P}_1 \oplus \mathcal{P}_2$$

durch das g-tupel $(P_1^3, \cdots, P_g^3)$ gegeben, das so bestimmt wird, daß es eine Funktion h auf C gibt, deren Nullstellen (mit Vielfachheit) genau gleich $P_1^1, \cdots, P_g^1, P_1^2, \cdots, P_g^2$ und deren Polstellen gleich $P_1^3, \cdots, P_g^3$ und P_∞ (mit Vielfachheit g) sind.

Die Existenz von h bzw. $P_1^3, \cdots, P_g^3$ wird durch den Satz von Riemann-Roch gesichert, die Berechnung von h wird durch Interpolationsalgorithmen möglich.

Allerdings sind diese Algorithmen im allgemeinen sehr zeit- und speicherplatzaufwendig, so daß wir bisher erst von einer befriedigenden theoretischen Lösung der ersten der uns gestellten Aufgaben sprechen können. Wir werden nun sehen, daß für eine spezielle Klasse von Kurven auch praktisch anzuwendende Algorithmen zu finden sind.

Betrachten wir das einfachste Beispiel: $g = 1$.
Da wir die Existenz eines K-rationalen Punktes vorausgesetzt haben, folgt wieder aus dem Satz von Riemann-Roch,daß C gegeben werden kann durch

$$Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

mit Koeffizienten $a_i \in K$ und der zusätzlichen Eigenschaft, daß kein Punkt von C singulär ist.
Kurven, die in dieser Form beschreibbar sind, heißen *Elliptische Kurven.* Daß solche Kurven von großem theoretischen Interesse sind, wird nicht nur mit Wiles' Beweis der Fermatvermutung belegt; hier interessieren wir uns also für ganz praktische Anwendungen ihrer Theorie.

Ein wesentlicher Grund für die Reichhaltigkeit der Theorie der elliptischen Kurven ist, daß sie gleich ihrer Jacobischen Varietät und somit gerade die Abelschen Varietäten der Dimension 1 sind. Insbesonders kann man, wie oben beschrieben, die Menge der rationalen Punkte zu einer Gruppe machen: Wir wählen den Punkt mit den homogenen Koordinaten $(0,1,0)$ als P_∞. Tatsächlich ist er der einzige Punkt auf C mit $z = 0$. Alle anderen Punkte P von C können mit affinen Koordinaten (x,y), die die Gleichung

$$(*)\quad Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$$

erfüllen, beschrieben werden.

Mit dieser Normierung ist

$$C(K) = \{(x,y) \in K \times K;\ (*) \text{ ist erfüllt}\} \cup \{P_\infty\}.$$

Wie sieht die Addition explizit aus? Jedenfalls entspricht P_∞ (wegen $P + P_\infty - 2P_\infty = P - P_\infty$) dem Neutralelement der Addition. Wir wollen nun für beliebige $P_1, P_2, \in C(L)$ die Koordinaten des Punktes angeben, der der Summe von P_1 und P_2 entspricht.

Um die Formeln zu vereinfachen, nehmen wir an, daß die Charakteristik von K ungleich 2 oder 3 ist (d.h. man darf in K durch 2 und 3 "dividieren") und vereinfachen (mit quadratischer Ergänzung bzw. Tschirnhausen-Transformation) die Gleichung für C zu

$$(**)\quad Y^2 = X^3 - g_2X - g_3$$

("kurze" Weierstraß-Form).

Wir betrachten nun Punkte $P_1 = (x_1, y_1)$ bzw. $P_2 = (x_2, y_2)$, die ungleich P_∞ sind und haben die Aufgabe, einen Punkt P_3 auf C so zu finden, daß es eine

Funktion h auf C gibt, die Polstellen der Ordnung 1 in P_3 und P_∞ hat und Nullstellen der Ordnung 1 in P_1 und P_2. Man kann diese Funktion, unter Benutzung von $(**)$ in der Form $h_1(X) + h_2(X)Y$, wobei $h_i(X)$ rationale Funktionen in X sind, ansetzen und erhält:

Falls $x_1 \neq x_2$ ist (allgemeiner Fall), ist

$$x_3 = -(x_1 + x_2) + \frac{(y_1 - y_2)^2}{(x_1 - x_2)^2}$$

und y_3 so, daß (x_3, y_3) auf C liegt und daß (x_1, y_1), (x_2, y_2) und $(x_3, -y_3)$ kollinear sind.
Falls $x_1 = x_2$ und $y_1 = -y_2$ ist, ist $P_3 = P_\infty$, und falls $x_1 = x_2$ und $y_1 = y_2$ ist (also $P_1 = P_2$) ist, ist

$$P_1 + P_2 = 2 \cdot P_1 = \left(\frac{X^4 + 2g_2X^2 + 8g_3X - {g_2}^2}{4X^3 - 4g_2X - 4g_3}, y_3\right)$$

(Verdoppelungsformel), wobei y_3 jetzt so gewählt ist, daß $(x_3, -y_3)$ auf der Tangente durch P_1 an C liegt. Man sieht, wie explizit und einfach die Addition auf C gegeben ist, und deshalb verwundert es nicht, daß die Vielfachenbildung sehr schnell durchführbar ist. Über Körpern K mit ungefähr 10^{50} Elementen und Zahlen $k \approx 10^{50}$ dauert auf einem gebräuchlichen PC die Berechnung von $k \cdot P$ bei sorgfältiger Implementation ungefähr 100 ms.

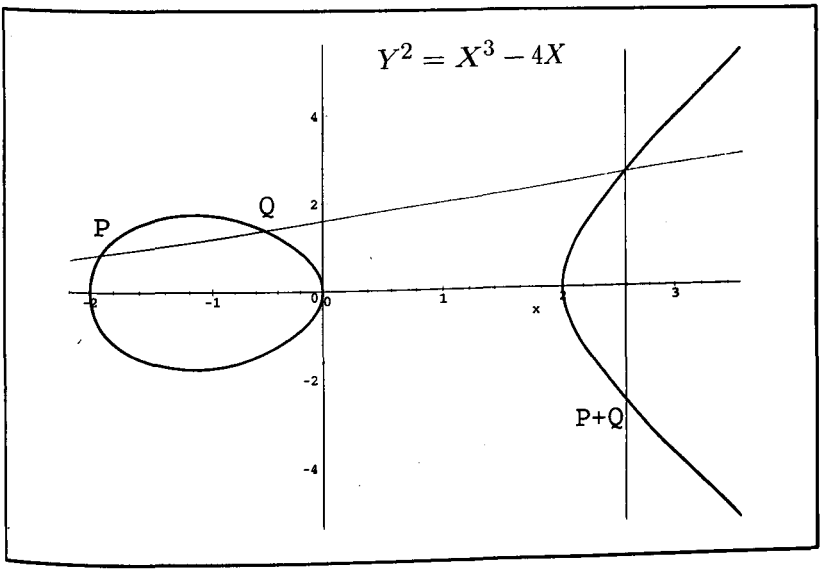

Abb. 1

Die Addition auf elliptischen Kurven hat noch eine sehr hübsche elementargeometrische Beschreibung: $P_3 = P_1 \oplus P_2$ ist der Punkt auf C, der durch Spiegelung an der X-Achse des dritten Schnittpunkts der Gerade durch P_1 und P_2 (resp. der Tangente in P_1, falls $P_1 = P_2$ ist) entsteht.
Wir können also feststellen, daß für Kurven vom Geschlecht 1 die erste der oben gestellten Aufgaben voll zufriedenstellend gelöst ist.

Bei Kurven C mit höherem Geschlecht beschränken wir uns auf solche, die durch ähnlich einfache Gleichungen wie elliptische Kurven gegeben werden: Eine affine Gleichung für C sei

$$Y^2 + a_1(X)Y + a_2(X) = f_n(X),$$

wobei $a_1(X), a_2(X)$ und $f_n(X)$ Polynome mit Koeffizienten aus K sind und $f_n(X)$ keine vielfachen Nullstellen besitzt. Eine milde Rationalitätsvoraussetzung (K-Rationalität eines Weierstraßpunktes von C) gestattet es, die Gleichung von C so zu wählen, daß wir genau einen Punkt P_∞ von C nicht durch diese affine Gleichung erfassen und daß die zugehörige ebene affine Kurve singularitätenfrei ist. Es folgt, daß $n = 2g + 1$ ist. (g ist wie immer das Geschlecht von C.)
Die einfache Gleichung für C gestattet es, die Beschreibung der Punkte $\mathcal{P}$ von $J_C(K)$ mit Hilfe von Idealen $I_\mathcal{P}$ des Rings $K[X] \oplus K[X]Y$ ($K[X]$ ist der Ring der Polynome in X mit Koeffizienten zu K, und Y hat die Eigenschaft, daß $Y^2 = f_n(X)$ ist) zu beschreiben, die eine Idealbasis der Form $(1, f_1(X) + f_2(X)Y)$ mit $f_i(X) \in K[X]$, Grad $(f_1(X)) <$ Grad $(f_2(X)) \leq g$ besitzen. Man kann also $\mathcal{P}$ durch (f_1, f_2) und damit (nach Normierung von f_1) durch die $2g$ Koeffizienten von f_1 und f_2 beschreiben. Ebenso ist die Addition beherrschbar: Zunächst multipliziert man die den Punkten $\mathcal{P}_1$ und $\mathcal{P}_2$ entsprechenden Ideale und reduziert dann, modulo Hauptidealen, das entstehende Ideal zu einem Ideal vom "Grad $\leq g$."

Eine Beobachtung von E. Artin in seiner Dissertation macht diese Rechenschritte besonders durchsichtig: Den Idealen kann man binäre quadratische Formen über $K[X]$ (mit Diskriminante $f_n(X)$) zuordnen und die Gruppenverknüpfung auf $J_C(K)$ entspricht der Komposition und Reduktion solcher Formen, die ganz entsprechend der Komposition und Reduktion definiter Formen über $\mathbb{Z}$, die Gauß hergeleitet hat, ablaufen. Der so entstehende *Algorithmus* zur Addition auf $J_C(K)$ für hyperelliptische Kurven C wurde von D. Cantor entwickelt, er verwendet als wesentliches Hilfsmittel nur die Division mit Rest von Polynomen. Im Prinzip kann man ihn auch dazu verwenden, *Additionsformeln* wie im elliptischen Fall herzuleiten, für $g = 2$ ist dies in der Dissertation von A. Spallek (Essen 1995) durchgeführt. Für höheres Geschlecht werden diese Formeln aber so kompliziert, daß die algorithmische Implementation der Addition vorzuziehen ist. Sie wurde in Essen durchgeführt und liefert bezüglich Einfachheit und Schnelligkeit mit den Ergebnissen bei

elliptischen Kurven konkurrenzfähige Resultate. Der Vorteil bei der Wahl von Kurven höheren Geschlechts zur Konstruktion von Falltürfunktionen liegt in der Möglichkeit, den Grundkörper K zu verkleinern, ohne die Sicherheitsanforderungen zurückzuschrauben. So kann etwa bei $g = 5$ bei Vorliegen einer 64-bit-Arithmetik auf Langzahlarithmetik verzichtet werden.

Wir müssen uns nun der zweiten Aufgabe, nämlich der *Konstruktion* geeigneter hyperelliptischer Kurven zuwenden. Ein naives "Ausprobieren" scheidet wegen der Größe von p aus. Die zu verwendenden Hilfsmittel kommen wieder aus der arithmetischen Geometrie. Die Galoisgruppe G_K endlicher Körper enthält ein ausgezeichnetes Element, den Frobeniusautomorphismus π_K, der, wie oben schon erwähnt, auf $J_C(\bar{K})$ operiert. Diese Operation wird nach Hasse-Weil dazu verwendet, ihm ein charakteristisches Polynom, die *L-Reihe* $L_C(T)$ von C zuzuordnen. $L_C(T)$ ist ein normiertes Polynom vom Grad $2g$, das als Analogon der Riemannschen Zetafunktion zu sehen ist. Die Nullstellen von $L_C(T)$ sind ganz-algebraische Zahlen, deren komplexer Betrag gleich $\sqrt{\#K}$ ist, dieser Satz (von Hasse für $g = 1$ und zuerst von Weil für beliebiges Geschlecht bewiesen) ist das Analogon der Riemannschen Vermutung. Das Polynom $L_C(T)$ sagt Wesentliches über C und J_C aus, unter anderem gilt:

$$L_C(1) = \#(J_C(K))$$

Mit anderen Worten: Die nun vorgegebene Konstruktionsaufgabe lautet jetzt: Man finde hyperelliptische Kurven über K, deren L-Reihe die Eigenschaft hat: Es ist $\left|\frac{L_C(1)}{p}\right|$ eine Zahl $\leq 10^5$.

Da die arithmetischen Eigenschaften der Nullstellen von $L_C(T)$ u.a. durch Arbeiten von Hasse, Weil, Deuring, Tate, Honda sehr gut bekannt sind, löst man diese Aufgabe, indem man erst geeignete Kandidaten für den *Endomorphismenring* $\mathcal{E}$ der gesuchten Kurve, in dem der π_K zugeordnete Endomorphismus liegt, sucht und dann die Kurve C konstruiert. Dieser Ring $\mathcal{E}$ ist eine Ordnung in einem "Körper vom CM-Typ", was für $g = 1$ einfach bedeutet, daß er im Ring der ganzen Zahlen eines imaginärquadratischen Körpers liegt. Es liegt daher nahe, zur Konstruktion von C die Theorie der "komplexen Multiplikation" und ihre Weiterentwicklung durch Taniyama-Shimura ("CM-Varietäten") zu verwenden. Dies haben wir konkret für $g = 1$ und $g = 2$ getan. Ein so erhaltenes Beispiel ist die Kurve C gegeben durch:

$$Y^2 = X^5 - 140X^2 - 240X^2 + 3810X + 6928.$$

C ist eine Kurve vom Geschlecht 2, über $\mathbb{Z}/(15394628755070098943) \cdot \mathbb{Z}$ hat J_C genau $4 \cdot 592486486457086864793418655053917441269 7$ Punkte. Wir können damit eine Falltürfunktion konstruieren.

Zur Konstruktion von Kurven vom Geschlecht ≥ 3 versuchen wir, zusätzlich zur CM-Theorie die Theorie der *Reellen Multiplikation* auszunutzen. Wir kommen damit zwangsläufig zu den Modulformen, zugehörigen Modulkurven $X_0(N)$ und ihren Jacobischen Varietäten $J_0(N)$, ,von denen ganz zu Anfang des Vortrags die Rede war, denn nach der (verallgemeinerten und noch unbewiesenen) Vermutung von Taniyama sollen Jacobische Varietäten, die reelle Multiplikation haben, als Faktoren von $J_0(N)$ auftreten.

Mit den von uns implementierten Algorithmen können wir solche Faktoren systematisch bestimmen und entscheiden, ob sie zu hyperelliptischen Kurven gehören.

Falls dies so ist, gelingt es mit Hilfe der Invariantentheorie für solche Kurven C, ihre Gleichung (über $\mathbb{Z}$) explizit zu bestimmen und wegen der Kongruenz von Eichler-Shimura die L-Reihe von C modulo l mit Hilfe des Heckeoperators zu l zu berechnen.

Wir geben ein Beispiel: Die Kurve C, gegeben durch

$$Y^2 = X^7 + 3X^6 + 2X^5 - X^4 - 2X^3 - 2X^2 - X - 1$$

ist eine Kurve vom Geschlecht 3, deren Jacobische Varietät ein Faktor von $J_0(284)$ ist. Modulo 1000040399 hat J_C $3^2 \cdot 17^2 \cdot 384534054770831874994067$ Punkte.

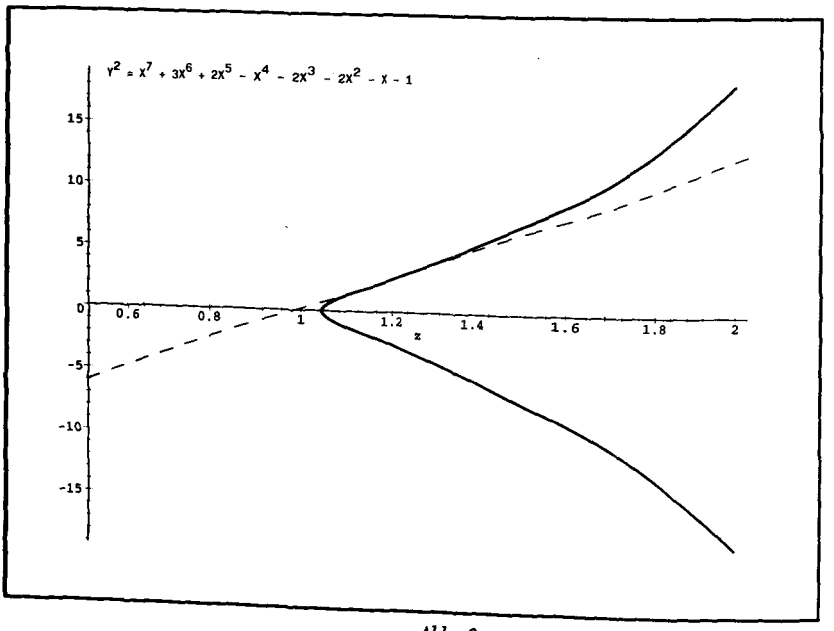

Abb. 2

Die geschilderten Konstruktionsverfahren benötigen einen hohen Rechenaufwand, z.B. die sehr präzise Auswertung komplexer Potenzreihen. Daß sie überhaupt durchführbar sind, verdanken wir der hochentwickelten Theorie, die uns von "reinsten" mathematischen Disziplinen, der Zahlentheorie und der Algebraischen Geometrie, zur Verfügung gestellt wird. Dieses Zusammenspiel zwischen mathematischer Grundlagenforschung und wissenschaftlichem Rechnen macht die Beschäftigung mit den geschilderten Aufgaben so faszinierend; daß uns durch die Anforderungen der digitalisierten Datenübertragung unverhofft Anwendungsmöglichkeiten eröffnet werden, ist eine hochwillkommene zusätzliche Motivation.

Gerhard Frey
Universität GHS Essen · Institut für Experimentelle Mathematik
Ellernstraße 29 · 45326 Essen